DOCUMENT RESUME

ED 121 331 IR 003 329

AUTHOR	Hall, Homer J.
TITLE	User Selection of Purchased Information Services. Interim Technical Report (June 1975-January 1976).
INSTITUTION	Exxon Research and Engineering Co., Linden, N.J.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	15 Mar 76
NOTE	16p.; For related documents see IR 003 327-328; Prepared by Government Research Laboratory
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS	Cost Effectiveness; Decision Making; *Information Needs; *Information Services; Interviews; *Library Material Selection; Research Libraries; Sciences; Special Libraries; Technology; Use Studies

ABSTRACT

Interviews conducted in the first phase of a project to develop a method for user selection of purchased scientific and technical information services identified a number of relationship among different populations of users. Research scientists, engineers, and patent attorneys want convenient access to original data identified in the search. Professional searchers select from a wider range of services. Their selection emphasizes the convenience of indexing methods or access to information available and the amount of training and time required to use each service to advantage. Information managers consider the budget questions of cost, time, and facilities and the feedback from users of all the services available. To this they add their own judgements of reputation, consistent standards of quality, and flexibility to meet the changing needs of specific users concerned. Also identified as a factor in the selection process was the rapid growth of scientific and information services which may be reaching a point where as many must shrink as can expand. (Author/PF)

ED121331

TECHNICAL REPORT

ON

USER SELECTION OF PURCHASED INFORMATION SERVICES

Contract No.: NSF C-1027

Project: Improved Dissemination and Use of Scientific and Technical Information

Interim Technical Report

June 1975 - January 1976

By

Homer J. Hall

Exxon Research and Engineering Company
Government Research Laboratories
P.O. Box 8
Linden, New Jersey 07036

March 15, 1976

IR003329

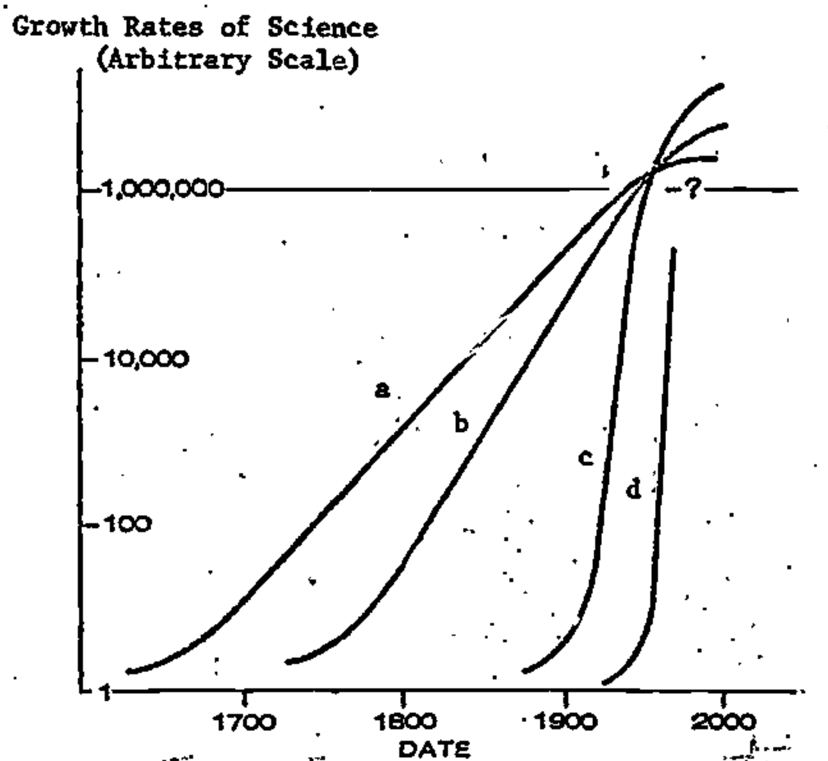
FACTORS IN THE SELECTION OF PURCHASED SERVICES

The growth of technology and science leads to new technologies, at an ever increasing rate. Each new generation in a given field of knowledge tends to grow faster, and takes less time to start. This concept of new generations in technology has some interesting corollaries, as proposed in 1960 by Derek de Solla Price. A number of these are summarized schematically in Figure 1. The initial rate of growth with time follows an exponential curve, a straight-line logarithmic plot, with a constant doubling period ( ). Exactly the same type of curve applies to the growth of money at compound interest, or bacteria in a culture, universities in European history, scientific journals in print, and theoretical returns from a chain letter. Price presents a number of illustrations from the history of science, and shows that the initial doubling rate for each of these is remarkably constant.

The fallacy of the chain letter applies to every growth curve which assumes continued exponential growth. Even the mathematics of compound interest assumes the permanence of an institution which can pay it, and a holder who is willing to wait. The hypothetical bank paying 6% a year since 1626 on a deposit of the $24 used to purchase Manhattan from the Indians should add $17 billion to the account in 1976, and $68 billion for the year 2000. But.... the whole deal is off if the bank goes broke, or if the impecunious holder decides to cash in his earnings at any time over the 350 years. There is always some resource which is consumed by growth, and as its disappearance becomes noticeable the exponential curve slows and flattens off to an S-shape, as shown. If the limiting resource can be replenished up to a point, like time or available manpower, this determines the level which the growth curve approaches as it flattens out. Mathematically, the S-shaped "sigmoid curve" contains an exponent which is diminished by a factor of $a(1-x)$, where x is the fraction of a vital resource which has been consumed. If there is no means by which it can be replenished, like bacteria with a limited food supply or a chain letter which has run out of gullible victims, growth goes negative, and the whole system collapses.

The second point Price makes is that in the history of science, a new technology is often developed and applied some time before the old has disappeared. This new technology can benefit from the methods and mistakes of the old, so it grows more rapidly, curve (b) in Figure 1. The next generation (c) grows still more rapidly and as it grows, the pressure which it creates for the development of succeeding generations can appear in a shorter and shorter time. Examples are given for the growth of scientific journals (plus abstract journals, and mechanical index systems); and for the overall growth of the scientific and technical literature in Europe, the United States, the Soviet Union, and China (up to 1960). Growth rates calculated for the last example correspond to progressively shorter doubling periods of 15 years, 10 years, 7 years, and 5 years respectively.

FIGURE 1



Schematic graph of the rise of science in various regions.

Growth Area	a	b	c	d
Scientific/technical population	Europe	U.S.A.	U.S.S.R.	China
Information services	Journals	Abstracts	Mechanical Indexes	Computerized Access

This regular pattern in the appearance of new generations of applied science suggests that there are some consistent factors involved. In each of these cases technology was available, waiting for a new need to develop: the printing press before the modern university, these and skilled makers of instruments before the scientific journals began, punched cards before mechanical searching, and so on. The pressure for the new development also involves a qualitative factor: 300 journals are more than any man can conveniently read, so the first abstract journals began to appear when the number of primary journals reached 300, and mechanical index systems began to appear when the list of abstract journals approached the same critical number. These can be represented by curves (a), (b) and (c) in Figure 1. Computerized systems for access to all of these are now appearing in a shorter time, and developing at a much more rapid rate (curve d).

This family of growth curves has another characteristic which is important and can be confusing, in that they overlap at the upper ends. Thus at the present time, computerized access to information (d) which started later than mechanical systems (c) and much later than abstracts (b) and primary journals (a) is growing so much more rapidly that it appears as a clearly dominant feature in the development of access to information. The NSF "guide to non-conventional scientific and technical information systems in current use" indicates that from 1958 to 1966 the number of mechanical sorting systems in use showed a doubling period of four years (e.g., edge punched, tabulating card, peek-a-boo systems). This rapid growth was all but obliterated by computer systems which increased at a sustained doubling rate of less than two years, from 17 in 1962 to 118 in 1966 (out of 175 non-conventional systems reported).

Overlapping Tools and Services

The purpose of the present study is to provide for the user of purchased services in scientific and technical information (STI) a guide to factors in the selection process.* The simultaneous development of different new methods of access to information confronts the user with an uncomfortable choice as to which tools to buy, and which buttons to push. It overlaps significant user-designed improvements in the older familiar systems, some of which he/she herself** may have helped create.

Science and technology have now reached the 4th or 5th generation of information access tools, with new tools coming on so rapidly that the lines between generations are blurred. A host of special services of every imaginable type are superimposed on journals which are still developing, along with abstract journals, mechanical indexes, and computer access to all of these. The design of new systems and services is itself a well-established discipline, so that the availability of new technology is less and less of a limit. The number of information "services" available is in the thousands and still growing, and each of them has been created to meet the challenge of specific user needs.

---

* NSF/Office of Science Information Services, Contract C-1027

** With due apologies for the English language, "he" will be used hereinafter to include "she" or "they"; "hers" is included with "his" or "their"; "himself" equals "herself" or "themselves", etc.

Four generations of access tools are identified in Table 1, and these basic types have many varieties. But a new service must sell to stay in business, and the user selection of services is a two-way street - the user selects from services available (or pays to build one himself), and only the services that users select remain available ( ).

The match of needs to users is critical, and it becomes more significant as the alternatives increase. If we consider the types of information and services proposed, a useful match appears between certain types of users and the types of access tools or services which they select, for certain types of need.

Users and Services

The proposition we will examine is that there are identifiable types of users with different types of needs who will differ in choosing one type of service over another, in STI or related fields, and that some of their bases of selection can be identified and explored. This match between user and type of service is as important an aspect of the selection process as the match between needs and services, which has been so widely discussed.

The literature on information is beset by semantic problems. The "user" we are considering in this study, the one who selects an information service, may be a scientist, a library searcher, a manager, or an information analyst. Except for the rare individual who does all his own library work, it is unusual for the same person to be both "user" and customer (the one who is asking for information); in most cases the "user" is collecting information for someone else. "Purchased" requires that the service be commercially available and therefore viable, on the market, as either a service to order or one available to many users. Our study is directed specifically to scientific and technical information, although similar principles may apply to other areas of knowledge with different parameters, e.g., for time or the nature of data.

The "typical user" populations listed in Table 1 are by no means mutually exclusive: any of these users can and do use all of the types of services described. The table is based on comments received in interviews with a number of users grouped to suggest relationships between specific types of information, services, users, and needs.

"Information services" from which any user may select fall into three major categories:

- access tools which the user can use to do his own search, as listed in Table 1;
- a service that does simple searching for known facts, which the user can pay for instead of finding the information himself;
- a service for information analysis, where the user pays someone else to both find and analyze the information available.

The distinction here is that a "simple search" is one where the question is pre-defined. This may be a lengthy process, but any change in the question once agreed upon is considered as a new search. It is also

Table 1

4 Generations of Information Services

	Type of Information	Type of Service	Typical User	Need Satisfied
(1)	Original data	journals (+ books)	Scientist, engineer	Details from originals
(2)	References	abstracts, indexes	Library searcher, patent attorney	Decisions often based on abstract or title alone.
(3)	Mechanical indexes	index to sources	Manager, planning staff	Relative value of sources (scientist know this in his field)
(4)	Optimization of search	interactive searching	Information analyst (any of the above)	Redefinition of search, to match needs and info available
(1,2)	Current awareness	newsletters, journal columnists, S.D.I.	Individual, for browsing or specific interests	Many variants, coverage selected by an editor or by the individual

only a search for facts, or references, not the judgment of facts. As a corollary of this, a commercial service which does simple searching may state that any searcher on the staff, with proper training, can accept an incoming question and come up with the same answer. The customer for a simple search may get all the original literature and references, or he may take selections from a report (bibliography) and request specific items. In either case he does his own evaluating, or arranges to have it done, as a separate step after the search is completed.

"Information analysis" introduces a new set of interactions which differ from the simple search: the user continually redefines the question by analyzing the information obtained during the course of the search. "What is...?" can be a simple search; "What if..." is quite different, and requires a different level of personal or staff ability and experience. This may involve conferring with the customer, or the user may do all the analysis and supply only a digested product. The same information services, even the most sophisticated ones, may be used either for a simple search or for information analysis: the difference lies in the user, and how the information tools and services are applied. In any case, the information analyst is applying ahigher level of skills than in simple searching, because the analyst must understand thoroughly the technical language of the literature and the customer's needs, as well as the inner workings of the searching procedures used. ( ).

Types of use can also be characterized by the depth of detail required for different searches, as suggested in Table 2. The "types" shown here are the familiar terms used in patent searches, which is one of the first STI areas where the art of searching was developed into commercial services. The same terms can be generalized for other uses: a state-of-the-art review has a similar function before starting a new project, whether this be in research, marketing, manufacture, or a patent application. This is for "how to do," whereas the novelty search is for "exactly what (not) to do." Both of these are frequently a simple search, where the subject is exactly defined the search can be completed as ordered. The customer receives and studies for best possible use the search he has paid for, and with the knowledge thus acquired he makes a new decision whether to order a further search. Infringment suits, contract negotiations or questions of liability are more likely to involve the information analysis approach, redefining both the question and the concept of needs, for the best use of the information available.

Dimensions in the Selection Process

Economic principles dictate that the user should purchase information services whenever he can buy information in the form that he needs and can use, for less than it would cost him to get the same information otherwise. This situation includes many hidden factors, such as the value of processing time and proprietary concerns. There are also strong negative values that are frequently overlooked, for information which consumes time and space because it is considered "useful" but never used. Values to any user involve aspects in the selection between services which are highly individual, such as the user's personal facilities, interests, and ability (or inability) to use certain types of information to advantage if it does become available.

Table 2

Depth of Detail

Type of Search	Use	Project Type	Characteristics
State of the Art	project planning, background	for research, marketing, or manufacture	overview, broad, moderate depth; what has been done and how.
Novelty	specific data, positive and negative	during research, marketing, manufacture	what works and doesn't work; what will appeal to the public; preferred methods
Contracts	define limitations, ranges, alternatives	for negotiations, infringement study	what will other managers pay for
Liability	govt. regulations, civil limits, caveats	all of the above (FDA, ERA, etc.)	evaluated technology; differences; best available, ultimate effects

The literature on the evaluation of information systems is voluminous, confusing, and often contradictory. Much of this literature reflects strong individual points of view, written by the purveyors of information services who emphasize their concern with matters of cost, or by system managers concerned with the value of time. They agree with each other on the importance of cost, where other users may not. There is a basic reason for this in that both are in a sense vendors, who must sell their time and services to someone else to stay in busines. Each author chooses carefully his own meanings for key words, and the concepts they identify, to give a new slant to old problems. These authors are often skilled linguists who draw subtle distinctions betwen the concepts they discuss, but they have only so many words to draw on, and the resultant confusion is inevitable. As a simple example, the literature on the "evaluation of information systems" has grown rapidly during the past 10 years ( ), but most of this evaluation has been aimed at improvements in the design and operation of single systems, not the comparative evaluation of different systems.

This literature started with a heavy emphasis on factors in cost ( ), and gradually changed over the years to include cost/effectiveness and cost/benefit factors ( ). These are progressively more difficult to quantify, involving subjective elements. Many authors have recognized and bemoaned the problems of evaluating complex variables containing both objective and subjective elements which cannot be exactly measured ( ). The total number of factors recognizable in the selection between services can be expanded almost at will: a recent checklist recognizes 144 of them, under 7 different headings. Many of these are closely linked, however, and are not independent variables ( ).

A dimensional approach to this complex system is recommended. The idea is to group variables which are so linked that they tend to move together in any change or comparison. Table 3 suggests that cost, time, personnel, and physical facilities can be considered as one such group, in the dimension of quantifiable factors. That is to say, any circumstance which has a major effect on cost, time, facilities, or staff requirements is likely to have some effect on other variables in the same group. A second dimension which is well recognized includes the whole gamut of qualitative factors that cannot be measured exactly but which can be handled by the technique of subjective ranking. The exact names given to different factors or relationships in this group are harder to handle than the fact that so many of them tend to be interrelated, and move together when any of them is significantly changed ( ).

The decision process is not based just on facts and personal preferences, however. Eloquent testimony on this can be offered by many a development engineer who is baffled when management has turned down a pet project that looked very good on paper. Management judgment tends to be a yes/no answer. This judgmental dimension is hard to argue with, based on neither numerical facts nor subjective rankings alone but a blend of these with intuition, based on experience ( ). It often involves a major contribution from the judgment of other managers. It is harder to define and less often discussed in the literature on information systems than either cost or cost/efficiencies/benefits, but perhaps this is only because it is so often taken for granted.

Table 3

Dimensions in the Selection Process:

Quantitative	Cost	Cost/efficiency, Cost/benefits
	Coverage	Fields, numbers, %
	Time for Entry	Text and index
Qualitative	Format	Appearance Convenient access
	Flexibility	Depth can be varied
	Responsiveness	Change priorities, add new items
	Feedback	To/from user
Business Judgment	Reputation	"Our way of doing things" Past recommendations Uniformity of quality
	Viability	Will it last
	Consistency	Dependable bias

The dimensional approach to the analysis of complex variables in project evaluation was advanced 15 years ago for the rational selection between projects in the management of industrial research. Harris (Monsanto) described a profile scheme with 5 "aspects" as headings for 26 factors, each ranked on a four-unit scale from very good to very bad ( ). This system for the evaluation of research was not widely adopted, and later studies suggested that 5 major headings is still too cumbersome. A strictly limited number of dimensions is preferred ( ). Three dimensions has this advantage, and is the minimum required to escape the hazard of false cause/effect conclusions from two-dimensional thinking.

The usefulness of the three dimensions proposed in Table 3 tends to be confirmed by the fact that their order of importance is quite different to different users. Initial comments received from different populations of users suggest the apparent relationships shown in Table 4. These are now being examined in further interviews:

- The scientist/engineer chooses first of all the service he finds most convenient to use. He accepts the judgment of a manager as to what systems to consider, and he pays little attention to the quantifiable factors of cost, time, or facilities as long as the systems available give him the answers he needs.

- The planning staff or information manager makes a first choice on whether the service in question is reliable, likely to stay in business, or "suits our way of doing things." He has to consider budgets on all the quantifiable factors, and may include qualitative factors only after he has ruled out services which he finds unreliable or too expensive.

- The literature on information systems emphasizes cost, time, and all the quantifiable factors a vendor must control to keep his customers. Qualitative factors are recognized in cost/benefit studies, but somewhat grudgingly because they cannot be accurately measured. Judgment as a separate dimension is well recognized in discussions on business management, but not usually in the literature on information services.

Initial interviews with a group of six searchers and scientists in the broad field of petroleum-related chemistry and industry provided a list of the information searching tools they consider most useful in their daily work, Table 5. This list showed a high degree of agreement, with minor differences in the order for specific individuals. No such agreement appeared in efforts to prepare a broader list of this type, to rank the 30 to 50 abstracting/indexing services the group considered most useful: titles and rankings beyond the few in Table 5 were completely different for each user/searcher, or for the same searcher for different searches. The next question was to prepare a trial list of discriminating factors which these searchers recognize as a basis for the selection between services. These factors expand the table into a grid, which is not shown here because it involves diverse statements of opinion on matters of fact. These differences in opinion are

Table 4

APPARENT SIGNIFICANCE TO GROUPS OF USERS

Dimensions (Groups of Factors)	Approximate Order of Significance	Selected Comments
Quantifiable, numeric (cost, time, facilities)	1 literature on systems 2 searcher, manager 3 scientist, attorney	vendors want all the data they can find balance time and costs against quality choose quality and dependability before costs
Qualitative, ranking (subjective, convenience)	1 scientist, searcher 2 literature 3 manager, planner	use what I like best, on list "approved" by others component of all cost/benefit analysis consider it only for services "reliable", on the market
Judgmental, yes/no (intuition, experiential)	1 manager, planner 2 scientist, searcher 3 literature	fits our way of doing business; will it last is it dependable de-emphasized; not "measurable"

Table 5

"Unique" Services, Most Used	Bases for Comparison
Chemical Abstracts	Fields covered
Index Hard Copy Text Microfilm	Cost $/yr Shelf space/yr
Derwent - C.P.I.	Timeliness: text + index
API - Abstracts:	User requirements
Patents and Literature	training skills
Science Citation Index	Minimum use required Maintenance time
Predicasts,	Comments
CMA, Funk-Scott	
Engineering Index	
Applied Science and Technology Index	

being explored as possible leads for new factors, or the linking of the factors named. Cost data were available from departmental staff, but they were not usually mentioned or even well known by the searchers.

The factors tabulated tend to be quantitative, and the comments qualitative or a complex of both. Factors in the dimension of judgment are applied to put or keep services on the list, but this decision is made at a different time and place. Once made, it is not a significant factor in the selection between services for daily work. Follow-up interviews indicate that users at other locations may be less interested in patents, or they may consult C.A. less and other services such as MEDLARS or BIOSIS more, but they apply much the same discriminating factors in selecting among them.

Status and Conclusions

A number of relationships can be identified among different populations of users. In general, research scientists and engineers agree that they want convenient access to original data identified in the search: the engineers want quick access to representative data, while the research scientists more often want details from all the different originals for comparison. The patent attorney also wants selected originals, but he is usually more interested in exact wording than in any data. Professional searchers select from a wider range of services, which they know how to use. Their selection emphasizes the convenience of indexing methods or access to information available, and the amount of training and time required to use each service to advantage. Patent attorneys and planning staff both usually have a well-defined question, and they emphasize wanting to deal with an individual searcher whom they know, who will understand their instructions and use mutually acceptable criteria in the selection of references. Information managers consider all the budget questions of cost, time and facilities, and the feedback from users of all the services available; they then add to this their own judgment of reputation, consistent standards of quality, and flexibility of each service to meet the changing needs of specific users concerned. This match between users and use is not random, and it is as important as any other aspect in the match between user needs and services.

Additional correlations of this type are being developed from interviews with a larger group of users in each of these populations. Special groups are also being considered, such as all the users and staff of a small branch industrial library. The present report covers the first six months of a study which started with user interviews in a major industrial research company, and is now branching out to other users in different disciplines and organizations of different size. A special search of the literature was completed but its analysis was deferred to permit the examination of reasons for some of the strong differences in points of view expressed. This search is now being checked against the conclusions outlined to date for additional data and differences.

The dimensions and factors proposed in Table 3 with respect to services in scientific and technical information have obvious analogies which may be useful in other areas. Comparable quantitative parameters which are limiting in biological systems include the total food supply (= total budget, or all the data available), total living space (= total volume or storage capacity), and time vs. metabolic rate (= time to get data, or time to digest data vs. total lifetime of the system). Parameters and problems in the relationship between costs and benefits are also being actively explored in many areas of pollution control ( ).

The rapid acceleration in the development of new technologies in STI may be approaching the point in the competition between systems where as many must shrink as can expand, and in the limit as many must die as can be born. This conclusion is not absolute, but there is a limiting parameter: the percentage of time available which is devoted to information services (or any other one area) can only be increased at the expense of other major areas, in research, development, or production.